图书情报工作
第62卷 第8期 2018年4月

知识共聚:领域分析视角下的知识聚合模式*

陈果[1] 吴微[1] 肖璐[2]
[1] 南京理工大学经济管理学院信息管理系 南京 210094 [2] 南京财经大学新闻学院 南京 210046

摘要:[**目的/意义**]当前知识聚合模式研究侧重"依据何种知识关联开展知识聚合",本文补充性地探索"利用知识关联将知识单元聚合成何种形式"这一后续问题,以完善知识聚合模式研究和引导实践的深入。[**方法/过程**]借鉴化学领域中对聚合反应两大类型划分的方法,提出以"是否保留知识单元间及其关联间的差异性"为根据,将知识聚合划分为知识类聚和知识共聚,并探讨知识共聚的基本实现形式。[**结果/结论**]领域知识是知识共聚开展的基础;以文档、词语为基本知识单元粒度,以用户需求入口和聚合目标资源为维度,知识共聚可通过四种基本形式实现:基于分面导航、基于多维概念关联推荐、基于知识元链接、基于资源潜在关联发现。

关键词:知识组织 知识聚合模式 知识共聚 领域知识分析 领域概念关联
分类号:G250
DOI:10.13266/j.issn.0252-3116.2018.08.015

1 引言

当前,知识聚合是图情学科的研究热点。"聚合"一词与英文"Aggregation"相对应,是指由异类成分组成的聚合,或多个细小单元组合形成的整体[1]。具体应用层面,较为典型的是化学领域中的"聚合反应",指单体小分子通过相互连接形成新的高分子材料[2]。在信息科学及其相关领域,以知识、数据、信息、资源为对象,又有"聚合""融合""集成/整合"等概念。李亚婷对相关概念进行了辨析和界定:聚合侧重于挖掘对象间多种关联并建立相应体系;融合强调对象经处理后产生的新变化;集成、整合则关注利用对象间的共性规则对其进行综合管理[3]。知识聚合是在信息资源快速增长导致有价值的知识资源高度分散和无序分布的背景下,通过发现"知识碎片"间的关联来对知识资源进行多维组合[4],以向用户提供体系化的、既可横向扩散又可纵向深入的知识内容。与数据集成、信息整合相比,知识聚合在聚合强度、聚合粒度上都有质的提升。知识聚合是实现基于用户需求的多元化、深层次、开放性知识组织和利用的前提。

近年来,国内图书情报领域以"聚合"为题的研究项目较多,因而也涌现出一大批成果。关于知识聚合项目和研究成果,大多以学术资源(以馆藏资源为典型代表)为对象[5-6]。而关于知识聚合模式的研究,大多侧重从聚合所依据的知识关联来区分。例如,贺德方等人将基于语义的馆藏资源深度聚合方式归纳为基于概念关联、引证关系和科研本体三类[5];赵蓉英等人将馆藏资源聚合模式归纳为传统聚合模式、深度聚合模式,前者包括基于编目和基于元数据的模式;后者包括基于本体、基于关联数据、基于主题模型和基于计量分析的知识聚合模式[7]。此后关于知识聚合的研究,也大都可纳入这些类型。值得一提的是,曹树金等人在梳理信息聚合的概念、构成后,提出的信息聚合模式不仅包括基于关系的聚合、粒度聚合,还包括基于情景的聚合[8]。

尽管当前研究已深入地探索了"应该依据什么来聚合知识",但这只是知识聚合模式的前半段问题。在获得丰富的、深层次的知识关联后,"利用这些关联将知识单元聚合成什么样",也应是知识聚合模式研究所解答的问题,但当前这方面研究相对较少。相当一部分知识聚合研究中,花费大量精力得到了深层次的知识关联(如语义关联),但最终它们被用于对知识单元

* 本文系国家社会科学青年基金项目"领域分析视角下的科技词汇语义挖掘与知识演化研究"(项目编号:16CTQ024)研究成果之一。
作者简介:陈果(ORCID:0000-0003-2873-1051),讲师,博士,通讯作者,E-mail:delphi1987@qq.com;吴微(ORCID:0000-0002-6111-0495),硕士研究生;肖璐(ORCID:0000-0001-5485-1407),讲师,博士。
收稿日期:2017-08-24 修回日期:2018-01-07 本文起止页码:115-122 本文责任编辑:王善军

进行聚类,例如:优化相似度计算以划分类团,或者构建关系网络以划分子群。

事实上,聚类只是知识聚合的一种初级方式。在充分利用知识关联细粒度、多样化特性的前提下,可针对用户需求和知识资源特点实现多种知识聚合方式。正如在化学领域,不同类型分子单体通过多种作用力聚合可形成丰富多样的共聚化合物。受此启发,本文从"如何充分利用知识关联的多样化差异开展知识聚合"为目标,提出知识共聚模式。首先,将知识聚合划分为知识类聚和知识共聚这两种基本模式并论述其特点;其次,从领域知识分析视角探讨实现知识共聚的基础;以此为基础,提出"词语 - 文档"体系下知识共聚的四种基本实现形式。

2 知识类聚与知识共聚

在化学领域,根据参与聚合的单体类型是否一致(也即聚合作用力是否单一),可将聚合反应分为均聚反应和共聚反应两种模式:均聚反应是指同类单体分子按同一作用力形成聚合物,而共聚反应是指多种类型的单体分子按不同作用力形成聚合物[2]。相应地,在知识聚合中也存在类似的模式差别:如果不考虑知识单元的类型差异,将其关联简化为单一的"相关关系"(典型的如相似度),以此为依据的聚合称为知识类聚;如果考虑知识单元的类型差异,以知识单元差异视角下形成的细粒度知识关联为依据,开展的聚合称为知识共聚。

2.1 知识类聚

知识类聚是最原始、直接且应用最普遍的知识聚合模式,由于在聚合中不考虑知识单元类型差异,因此"相关关系"的量化计算是核心。常见的知识单元相关关系有共现关系和相似性关系。共现关系是指知识单元在相同时间、空间或知识场景中出现,其量化方法有邻近指数、等价系数、基于改进的 TF-IDF 算法等。相似性关系的量化主要包括三类:一是以个体共现关系为基础进行的相似度计算;二是依据一定的语义知识库将个体的语义关联转化为相似度;三是融合前两种相似度结果,称为语义相似度计算[9]。

知识类聚的实现形式主要有两种:一种是将知识单元置于相应的特征空间内,以距离为依据来划分类团,典型的如词语或文本的聚类、分类;另一种是将知识单元及相关关系表示为网络,基于特定的网络指标和网络子群发现技术来划分类团,典型的如共词网络中词簇划分。由此可见,聚类只是知识类聚的一种实现方式。

马文峰等人指出,以分类法/主题词表为代表的概念类聚体系不强调对概念间关联的组织,尚未深入到语义层面,因此是较低层次的知识组织体系[10]。相应地,知识类聚也是较低层次的知识聚合模式。类聚结果将多维知识关联映射到单一维度的相关关系上,不利于实现知识资源的多维组合,因此在实践应用中存在明显的缺陷。其典型问题是"知其然而不知其所以然",即:知道类聚的结果间存在关联,但不知道存在何种关联,因而不利于用户需求的扩展、细化和引导。

2.2 知识共聚

知识共聚模式并非凭空而来。当前知识组织的主要任务已由排序和归属转变为知识关联和链接的构建[11],在知识组织体系中,本体、概念地图相对于分类法、主题词表等更为高级,其原因是它们保留了知识单元和知识关联的语义类型差异[10]。类似地,知识共聚注重在特定知识场景下,将知识单元以多样化的关联为作用力聚合成可解决特定问题或任务的知识共聚体,因此它是一种更为深层的知识聚合方式。

在实践中,知识共聚的典型应用是分面导航,当用户检索或点击某一个主题时,分面导航系统会展现相关结果与检索对象之间多维关联,这种模式在网络信息组织上具有明显优势[12]。以检索"手机"为例,百度的相关搜索词引导和京东的商品导航分别对应知识类聚模式和知识共聚模式,后者包含多个与手机属性相关的侧面。此外,除了以词语(概念)为粒度开展知识共聚,以句子、文档、用户为粒度的知识聚合均可由知识类聚模式深化到知识共聚模式。例如,唐晓波等人提出对关于某一事件的微博按句子逻辑关系进行聚合,由"起因"到"事件"并导向"结果"[13]。

由此可见,知识共聚的结果是一个包含多个侧面(由知识单元及其关联的类型差异决定)的立体结构,从不同侧面进行解读可获得不同的知识单元类团。用户在得到知识共聚结果后,可根据其特定需求选择利用相应侧面。知识类聚的结果尽管也可视作一个立体空间,但它缺少知识单元差异支撑下的侧面,因而只能按照距离划分类簇。因此知识共聚模式保障了聚合结果的完整性和应用灵活性,相比知识类聚模式更符合用户知识探索和知识利用的需要。

3 知识共聚的基础:领域概念关联

3.1 知识组织中的领域分析视角

在知识生产、交流和利用中,用户需求和知识内容

通常具有领域性，因此知识组织需要以领域知识为背景开展。领域知识（Domain Knowledge）与通用知识（如常识、逻辑规则和数学运算知识）相对应，是指特定领域内重要概念及它们之间的关联[14]。在信息服务和情报工作中，仅仅追求信息组织方式的优化而忽略知识的领域特性，会导致“重形式轻内容”的问题[15]。因此，情报学研究方法中产生了经典的“领域知识分析范式”，其核心思想是立足于特定知识领域来审视人类信息活动行为[16]。实际上，领域知识分析中知识元素的组织仍然依托于传统的通用知识组织手段，只是在其中引入一些特定的技术、方法，以使之更符合面向特定领域的知识组织和知识服务目标。

面向特定领域进行知识组织时，以领域知识分析范式指导的知识组织体系可取得更好的效果。M. K. Buckland 等人的研究表明，依据特定领域专业语言编制的索引相比综合索引可带来明显的检索效果优化[17]。领域知识构成的基础单元是概念，因此 B. C. Vickey 等人认为生成领域性专门词表是领域分析的一个重要目标[18]；而概念关联又是领域知识关联的本质层面，因此，领域知识分析的一个关键目标是针对特定领域识别其概念和概念间的细粒度语义关联，即建立领域概念关联体系[19]。

3.2 领域概念关联及其在知识共聚中的作用

知识单元间的关系包括等同关系、层次关系和相关关系，通用性的知识组织系统基于这些关系构建。例如，汉语主题词表中的用、代、属、分、参等关系与等同、层次、相关关系对应。但是，特定领域知识组织更关注具体化和差异化的“相关关系”，其原因是：不同领域对同一概念的众多特征各有侧重，例如，“水”这一概念，在化学、物理学、生物学、环境科学等领域所被关注的本质特征差异明显[20]。相应地，引入领域背景知识后，通用层面粗略的“相关关系”可被区分为丰富多样的细粒度概念关联，例如方法与应用关系、疾病与症状关系。领域概念关联相对于一般性的概念关联而言，有两个明显的特征：概念的内涵具有明确定义，且概念关联的类型是具体、细化和多样化的[19]。

作为知识关联本质层面的领域概念关联是实现细粒度、多维化知识共聚的基础。在知识共聚模式中，领域概念关联的作用可概括为：保障知识资源组织结构与用户知识需求结构间的一致性，即：已有资源中知识单元及其关联的识别和利用应以领域概念关联体系为参照，用户需求的转化、扩展和细化也在领域概念关联体系对应的空间内引导完成。

在现有的资源系统中，知识内容的组织结构与用户需求结构并不一致。用户知识需求往往围绕特定任务的解决而产生，这就决定了其目标知识具有一定的领域性，即其所需知识受任务对应领域内知识结构（如概念、概念关联）的约束。由于对目标知识的未知，用户需求又是潜在、模糊的，其知识搜寻行为是一个根据结果不断学习、优化和反馈的迭代过程。然而，现有的资源系统中，知识内容以“词语”为代表分散在不同文档中。缺少领域知识为参考时，不得不基于“相关性”来建立的知识关联，以推测用户需求转化的方向。但这种方式并不能有效地实现用户需求引导及其与知识资源的匹配。研究表明，缺乏领域知识指导时，概念关联的挖掘会存在较多无意义的结果，难以满足用户的特定需求[21]。

实质上，无论用户需求如何模糊，也无论知识资源如何分布，两者之间在领域知识背景下具有一致性。在领域知识场景下，知识资源中的知识单元以不同形式的关联共同构成解决特定问题的知识共聚体。领域概念关联体系中所包含的明确的、丰富的概念类型和细粒度概念关联，是揭示词语、文档等知识单元间多维知识关联的有效依据，由此对知识单元进行多维、细粒度聚合，可交互式、渐进式地引导用户，使其知识需求在领域知识空间下有针对性地横向扩展和纵向细化，以有效地获取和利用知识。

4 知识共聚的四种基本实现形式

知识聚合的表现形式与其要素紧密相关，曹树金等人将信息聚合要素归纳为“聚合对象粒度、情景、关系”，并以情景分析为基础，以聚合机制解释和聚合应用介绍为主要内容，对信息聚合的模式和相应的表现形式进行了梳理[8]。以此为理论基础，知识共聚的基本实现形式如下：

聚合对象粒度：以知识资源系统中的词语（如关键词、标签、主题词、概念术语、知识元等）和文档（如网页、短文本、学术文献等）这两种最基本的知识单元为聚合对象粒度；

聚合情景：以用户知识探索和利用过程中两类知识单元间的转换为聚合情景。例如，通过关键词获取文章，阅读文章后探索感兴趣术语相关知识等。总体而言，包括在从词语到词语、从词语到文档、从文档到词语、从文档到文档四种基本情景。

聚合关系：以领域概念关联和“文档 – 词语”的包含关系为聚合依据，领域概念关联主要用于构建词语

间的多维关联体系,“文档－词语”的包含关系用于将文档资源的粗粒度关系映射到词语层面的细粒度概念关联上,实现对知识资源的多维、多粒度聚合和对用户有针对性地引导。

相应地,以词语、文档为基本粒度,以词语、文档在用户知识探索和利用中的切换路径为情景,以领域概念关联和“文档－词语”包含关系为聚合依据,可将知识共聚的几种基本实现形式归纳如图 1 所示。该图中的四个象限表明用户从对应的需求入口(横轴)探索对应的聚合目标(纵轴)时,可以利用到的知识共聚形式。

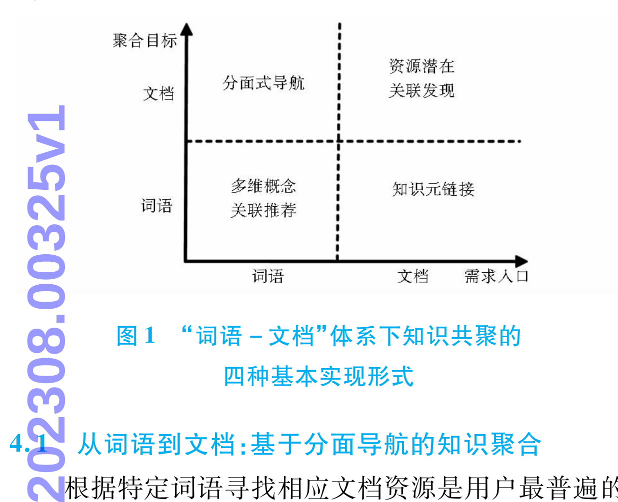


图 1　“词语－文档”体系下知识共聚的四种基本实现形式

4.1　从词语到文档:基于分面导航的知识聚合

根据特定词语寻找相应文档资源是用户最普遍的知识探索路径,按关键词检索文档、按导航词浏览文档都属于这种形式,可将从词语到资源(主要是文档)的引导统称为“导航”。由此看来,导航是知识资源聚合必然需要考虑的重要实现形式。

传统知识服务平台中,普遍存在导航结构扁平化、导航词间语义关联欠缺等问题,难以满足用户持续性、渐进式的知识探索需求[22]。因此,不少研究者从知识聚合的视角探索导航优化。李亚婷认为基于知识聚合的导航服务不应仅仅依赖于严格单一的层次结构,而应注重从内容中提取聚合类别[3]。胡媛等人提出了基于知识聚合的数字图书馆社区知识导航体系框架[23]。张云中等人以专家分类法与大众分类法互补融合为思路,提出了 tax-folk 混合导航模型,实现了树状“干强枝繁叶茂”的资源聚合[24]。在充分利用知识关联结构和资源属性特征的前提下,可构建分面式导航体系[25]。相对传于统基于相似度或分类/主题表的导航体系而言,分面式导航体系充分保留了结果文档与导航词汇的多样性关系,同时也更符合用户资源导航习惯。这也是知识共聚模式下的资源导航实现形式。

引入领域概念关联体系后,可对知识系统中原有的导航体系进行分面化改造。具体思路如图 2 所示:首先,从原有导航词中筛选出领域概念术语,以排除资源定位效果差的导航词、精简导航方向(如图 2 中,剔除了原导航体系的 b、f、g 等词);其次,基于领域概念类型和细粒度关联设置导航分面(如图 2 中所设的 3 个分面)。导航词各代表一个文档集合,由此将已有的大量文档资源按分面导航方式进行聚合组织。基于该思路,笔者以丁香园医学社区中的心血管论坛为例,构建了分面导航体系以实现 UGC 资源的知识共聚,其系统原型如图 3 所示[26]。

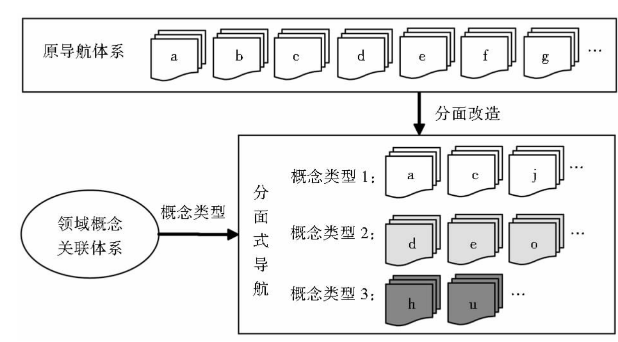


图 2　基于分面式导航的知识聚合实现形式



图 3　基于分面式导航的心血管领域 UGC 知识聚合示例[26]

4.2　从词语到词语:基于多维概念关联推荐的知识聚合

当用户检索或浏览某一词语相关文档时,与该词语相关的其他词语构成其目标知识的一部分。用户知识搜寻和利用行为是一个持续过程,通常需要不断地横向拓展和纵向深入。搜索列表页的查询词推荐、结果内容页的标签推荐等都属于“从词语到词语”的知识引导。因此,知识聚合的有效实现需要充分考虑这一场景。

传统的词语推荐通常依托日志、文档等语料,对词语按相似度计算和排序[27],再以线性方式列举推荐,其语义模糊性导致用户知其有关而不知其何关,不利于用户需求的进一步扩展和细化。因此,有必要加入语义信息以使之更符合用户知识结构。陆伟等人通过

陈果，吴微，肖璐. 知识共聚:领域分析视角下的知识聚合模式[J]. 图书情报工作,2018,62(8):115－122.

实验表明，采用主题分析可以显著提升查询推荐的精确度[28]。洪婕等人认为，在领域型知识系统中，依托语义资源库来构建查询推荐体系是一个很好的选择，并通过实验论证了利用领域本体可获得比百度、google查询推荐更好的效果[28]。靳延安则系统地讨论了基于语义粒度、话题敏感性、用户动机的标签推荐技术[30]。事实上，相关词推荐的核心目标是构造用户后续知识场景(如纵向深入、横向扩展)变化的空间。而词语的内涵通常基于它与其他词语的概念关系来表现，与当前词相关的其他词语极有可能是用户接下来需求扩展或细化的目标。因此，对相关词语按概念关联进行多维聚合以开展相关推荐，是实现"从词语到词语"场景下知识共聚的一种重要方式。

如图4所示，引入领域概念关联体系后，将用户入口词相关的概念及其细粒度关联引入到词语推荐中，形成具有多维结构的相关词推荐。一方面可排除与领域知识无关的干扰词；另一方面，通过标明相关词语与入口词的细粒度关联类型(如图4右侧a词的相关词具有不同推荐维度)，用户可以有针对性地选择下一步知识搜寻的方向。由此可实现对知识系统中词语的有效聚合。基于该思路，笔者以丁香园医学社区中的心血管论坛为例，构建了基于概念关联的多维推荐体系以实现UGC资源的知识共聚，其系统原型如图5所示。

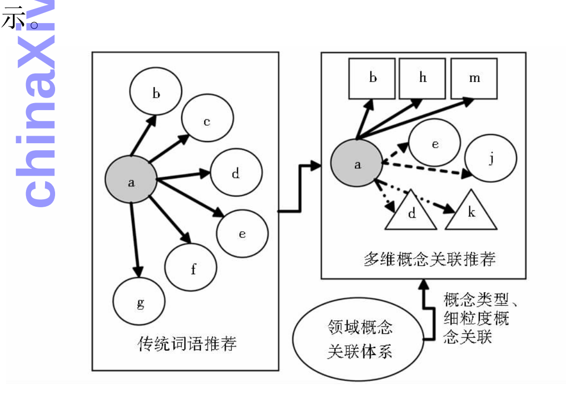


图4　基于多维概念关联推荐的知识聚合

4.3　从文档到词语:基于知识元链接的知识聚合

文档中具有语义内涵的概念单元可称为知识元，一篇文档的知识价值体现为将多个知识元按若干关联组织起来。用户浏览特定文档时，通常会对其中某些知识元感兴趣。如何以文档中的知识元为载体开展相关内容聚合，是解决知识在文档中"碎片化"分布、实现海量文本聚合向精准化知识聚合的重要途径。

知识元链接体系是在用户文档浏览场景下实施知识共聚的重要形式。曾建勋认为基于知识元链接构建



图5　基于多维概念关联推荐的心血管领域UGC知识聚合示例

知识网络是知识链接的一个重要方向[31]。孙震等人提出基于知识元的新型科学计量范式，其核心是以专业问题和学科的知识点作为计量口径和知识基础单元[32]。毕崇武等人认为将知识元以链接的方式关联起来，可以将最细粒度的知识元聚类为不同粒度的知识集合，并提出了基于知识元的多粒度知识集合组织方式[33]。陈果等人提出了融合领域知识库和共现分析的网络社区知识元链接体系构建方案，通过构建知识元链接体系，实现网络社区中碎片化知识的深度连通[34]。

基于知识元链接的知识共聚实现思路见图6。首先，通过参考领域概念关联体系中的概念术语对文档中的领域概念进行知识元标注，将其指引到相应的知识元内容页；其次，构建知识元内容页，其组成一般包括知识元的基本内涵、与该知识元相关的其他知识元、与该知识元相关的文档资源等。通过构建独立的知识元链接体系，可在不改变原有资源组织架构的前提下，在文档内实现相关知识元的内容聚合。基于该思路，笔者以丁香园医学社区中的心血管论坛为例，构建了知识元链接体系以实现知识共聚[34]，其文档中知识元标注结果示例见图7，知识元内容页示例见图8。

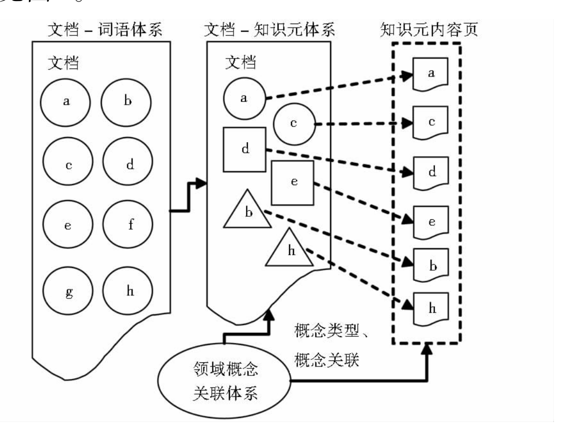


图6　基于知识元链接的知识聚合

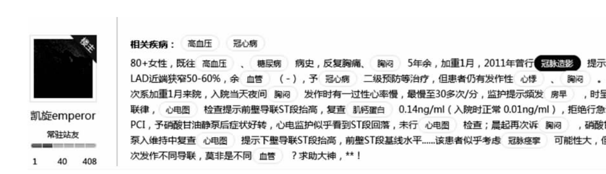


图 7　心血管领域 UGC 资源的知识元标注示例



图 8　心血管领域知识元内容页中的知识聚合示例

4.4　从文档到文档:基于资源潜在关联发现的知识聚合

文档是用户获取知识的最终载体,从文档到文档是知识系统中最直接的聚合形式。常见文本聚合形式是以其外部属性信息或内容标注为特征计算相似度,再进行文本聚类、分类等工作。这种聚合方式的主要问题是结果冗余度过高,而忽略了文档间的潜在语义关联,因而在解决面向任务的知识需求时存在明显缺陷。例如,用户浏览了关于“高血压”介绍的文档后,可能需要了解与其中某些症状相关的内容,而非更多关于“高血压”介绍的相似文档。

为解决这一问题,研究者从语义层面开展文档资源关联发现研究。陈兰杰和侯鹏娟将数字文献资源关联关系揭示方法划分为挖掘关联关系方法和构建关联关系方法,后者包括语义与本体方法、关联数据方法等,并指出各自优缺点[35]。赵夷平和毕强提出利用潜在语义分析与向量空间模型计算文档内容的相似性,再将语义关联信息和文档元数据合并为关联数据,实现后续的相似文献发现;他们同时指出,以客观知识体系和知识结构为基础的文献关联发现应更能体现相关文献的关联程度[36]。洪韵佳和许鑫提出了一种基于领域本体的适用于知识库树状结构的多层次文本聚类方法,实现了从粗粒度到细粒度的多层次聚类[37]。

在知识共聚模式下,以领域概念关联为依托,可参考非相关文献的知识发现的思路来建立资源关联[38]。即:在两个相对独立但其对应概念存在关联的文献间建立关联,进一步实现文档聚合。基本实现方法如图 9 所示:如果文档 A 中出现领域概念 a1,文档 B 中出现领域概念 a2,而在领域知识体系中 a1 和 a2 存在某种细粒度关系,则可在 A 与 B 之间初步标注这种关联关系;通过对 A 与 B 间更多初步关联的统计可确立其最终关联类型。在具体实践中,笔者针对丁香园医学社区中的心血管论坛开展了基于关联发现的文档聚合实验,以“冠脉造影”为例,只有 61% 的讨论贴是通过直接关联与“冠心病”讨论实现聚合,剩下 39% 需要依托领域概念关联体系冠脉造影与冠心病的诊断关系来实现聚合,其效果改进可见一斑[19]。

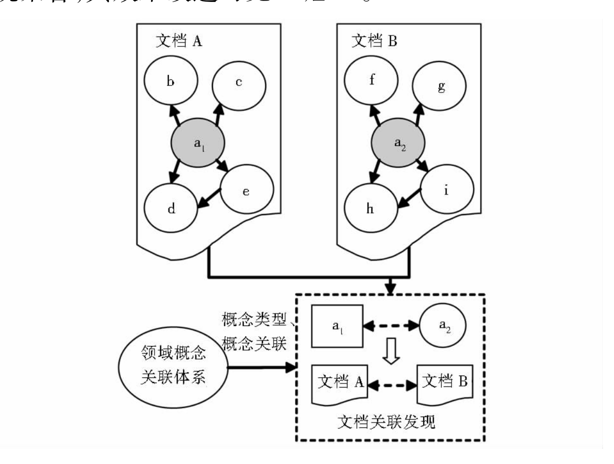


图 9　基于资源潜在关联发现的知识聚合

5　结语

知识聚合模式的研究不仅应关注基于何种知识关联来开展聚合,还应探讨利用这些知识关联将知识单元聚合成什么形式。鉴于后者相关研究较为缺乏,本文参考化学领域中聚合反应的两种类型,将知识聚合模式划分为知识类聚和知识共聚两大类,其区别为在聚合中是否保留知识单元及其关联的多样性。基于领域知识分析视角,可知知识共聚模式实现的基础是领域概念及其细粒度关联的有效利用。在“文档 - 词语”形式的知识资源体系下,以文档、词语为基本知识单元粒度,以用户需求入口和聚合目标资源为维度,可将知识共聚模式的实现形式划分为基于分面导航的知识聚合、基于多维概念关联推荐的知识聚合、基于知识元链接的知识聚合、基于资源潜在关联发现的知识聚合。

知识聚合的开展同时受资源组织结构和用户需求结构的影响,由领域概念及其细粒度关联构成的领域知识是保障两者一致性的关键。当前利用特定领域的知识背景以开展知识聚合的研究和实践较为缺乏,而针对特定领域资源开展的知识聚合与通用性的知识聚合在模式上存在很大区别。随着知识聚合研究和实践

的逐步深入,针对特定领域进行知识聚合将面临新的问题,本文所提出的知识共聚模式及其实现形式是一种基本途径,在后续研究和实践中也将得到进一步完善。

参考文献:

[1] THOMPSON D. 牛津现代英汉双解词典[M]. 北京:外语教学与研究出版社, 2003: 38.

[2] CLAYDEN J, GREEVES N, WARREN S G. Organic chemistry [M]. Oxford, UK: Oxford University Press, 2001.

[3] 李亚婷. 知识聚合研究述评[J]. 图书情报工作, 2016, 60(21):128 - 136.

[4] 赵雪芹. 知识聚合与服务研究现状及未来研究建议[J]. 情报理论与实践, 2015, 38(2): 132 - 135.

[5] 贺德方, 曾建勋. 基于语义的馆藏资源深度聚合研究[J]. 中国图书馆学报, 2012, 38(4):79 - 87.

[6] 毕强,尹长余,滕广青,等. 数字资源聚合的理论基础及其方法体系建构[J]. 情报科学, 2015(1):9 - 14.

[7] 赵蓉英, 王嵩, 董克. 国内馆藏资源聚合模式研究综述[J]. 图书情报工作, 2014, 58(18):138 - 143.

[8] 曹树金, 马翠嫦. 信息聚合概念的构成与聚合模式研究[J]. 中国图书馆学报, 2016, 42(3):4 - 19.

[9] 刘萍, 陈烨. 词汇相似度研究进展综述[J]. 现代图书情报技术, 2012, 28(7): 82 - 89.

[10] 马文峰,杜小勇. 关于知识组织体系的若干理论问题[J]. 中国图书馆学报, 2007(2): 13 - 17.

[11] 滕广青,毕强. 知识组织体系的演进路径及相关研究的发展趋势探析[J]. 中国图书馆学报, 2010(5): 49 - 53.

[12] 张振森, 程灏, 李丽. 网络信息分面分类体系的理论与应用研究[J]. 情报杂志, 2009, 28(8):62 - 66.

[13] 唐晓波, 肖璐. 基于单句粒度的微博主题挖掘研究[J]. 情报学报, 2014, 33(6): 623 - 632.

[14] 景丽萍, 恽佳丽, 于剑. 领域知识在文本聚类应用中的机遇和挑战[J]. 计算机工程与科学, 2010, 32(6): 88 - 91.

[15] 王琳. 领域分析范式视角下知识组织中若干问题研究[J]. 图书情报工作, 2011, 55(4): 90 - 105.

[16] HJφRLAND B, ALBRECHTSEN H. Toward a new horizon in information science: domain analysis[J]. Journal of the Association for Information Science and Technology, 1995, 46(6): 400 - 425.

[17] BUCKLAND M K, CHEN A, GEBBIE M, et al. Variation by subdomain in indexes to knowledge organization systems[C]//BEGHTOL C, HOWARTH L C, WILLAMSON N J. Dynamism and stability in knowledge organization: proceedings of the 6th international ISKO conference. Wuerzburg: Ergon Verlag, 2000:48 - 53.

[18] VICKERY B C. Faceted classification: a guide to construction and use of special schemes[J]. Library quarterly information community policy, 1960(3):289.

[19] 陈果. 基于领域概念关联的网络社区知识聚合研究[D]. 武汉: 武汉大学, 2015.

[20] 国家质量技术监督局. 术语工作原则与方法:GB/T10112 - 1999 [S]. 北京:中国标准出版社,1999:1 - 6.

[21] 梁凯强, 陆菊康. 基于领域本体与概念格的关联规则挖掘[J]. 计算机工程与设计, 2007, 28(13): 3033 - 3035.

[22] 毕强, 周姗姗, 马志强,等. 面向知识关联的标签云优化机理研究[J]. 现代图书情报技术, 2014(5):33 - 40.

[23] 胡媛, 陈琳, 艾文华. 基于知识聚合的数字图书馆社区集成推送服务组织[J]. 图书馆学研究, 2017(19):9 - 17.

[24] 张云中, 杨萌. Tax-folk 混合导航:社会化标注系统资源聚合的新模型[J]. 中国图书馆学报, 2014(3):78 - 89.

[25] HEARST M A. Clustering versus faceted categories for information exploration[J]. Communications of the ACM, 2006, 49(4):59 - 61.

[26] 陈果,肖璐,孙建军. 面向网络社区的分面式导航体系构建——以丁香园心血管论坛为例[J]. 情报理论与实践, 2017, 40(10):112 - 116.

[27] 吴思竹. 社会标注系统中标签推荐方法研究进展[J]. 图书馆杂志, 2010, 29(3):48 - 52.

[28] 陆伟, 张晓娟. 基于主题与用户偏好分析的查询推荐研究[J]. 情报学报, 2012, 31(12):1252 - 1258.

[29] 洪婕, 张健, 胡亮. 基于领域本体知识库的专业搜索引擎查询推荐算法研究——以盐湖化工领域为例[J]. 情报学报, 2014(10):1091 - 1098.

[30] 靳延安. 社会标签推荐技术与方法研究[D]. 武汉:华中科技大学, 2011.

[31] 曾建勋. 知识链接的研究现状与发展趋势[J]. 情报理论与实践, 2011(2): 119 - 123.

[32] 孙震, 冷伏海. 基于知识元的新型科学计量范式探析[J]. 情报学报, 2017, 36(6): 555 - 564.

[33] 毕崇武, 王忠义, 宋红文. 基于知识元的数字图书馆多粒度集成知识服务研究[J]. 图书情报工作, 2017, 61(4): 115 - 122.

[34] 陈果, 肖璐. 网络社区中的知识元链接体系构建研究[J]. 数据分析与知识发现, 2017, 1(11): 75 - 83.

[35] 陈兰杰, 侯鹏娟. 数字文献资源关联关系揭示方法研究[J]. 图书馆, 2015(2):41 - 45.

[36] 赵夷平, 毕强. 关联数据在学术资源网相似文献发现中的应用研究[J]. 现代图书情报技术, 2016, 32(3):41 - 49.

[37] 洪韵佳, 许鑫. 基于领域本体的知识库多层次文本聚类研究——以中华烹饪文化知识库为例[J]. 现代图书情报技术, 2014(12): 19 - 26.

[38] SWANSON D R. Undiscovered public knowledge[J]. Library quarterly, 1986, 56(2):103 - 118.

作者贡献说明:

陈果:提出思路,完成实验;

吴微:撰写初稿;

肖璐:查阅资料,修改论文。

Knowledge Copolymerization: A Knowledge Aggregation Mode Under the Perspective of Domain Analytic Paradigm

Chen Guo[1] Wu Wei[1] Xiao Lu[2]
[1] Department of Information Management, Nanjing University of Science and Technology, Nanjing 210094
[2] School of Journalism, Nanjing University Of Finance & Economics, Nanjing 210046

Abstract: [**Purpose/significance**] Current research on knowledge aggregation mode focuses on the types of knowledge connection that knowledge aggregation is based on. As an indispensable complementary, this paper explores the follow-up question of what form knowledge elements can be aggregated based on those connections, it aims at perfect the research of knowledge aggregation mode and guide related practice. [**Method/process**] Taking the classification of polymerization in chemistry, this paper put forward to divide knowledge aggregation mode into homopolymerization and copolymerization according to "whether or not remain the differences between knowledge units and their correlations", and then explored the basic realization forms of knowledge copolymerization. [**Result/conclusion**] Domain knowledge is the basis of knowledge copolymerization. Utilizing documents and words as two basic knowledge elements, and considering users' demand portals and target knowledge elements as two dimensions, knowledge copolymerization can be achieved through the following four basic forms: based on faceted navigation, multidimensional recommendation based on conceptual relation, based on knowledge element linking, and knowledge discover based on the potential connection of resources.

Keywords: knowledge organization knowledge aggregation mode knowledge copolymerization domain analytic paradigm domain conceptual relation